



OECD Economics Department Working Papers No. 1022

Educational Attainment
and Labour Market
Outcomes in South Africa,
1994-2010

**Nicola Branson,
Murray Leibbrandt**

<https://dx.doi.org/10.1787/5k4c0vvbv0q-en>

Unclassified

ECO/WKP(2013)14

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

19-Feb-2013

English - Or. English

ECONOMICS DEPARTMENT

ECO/WKP(2013)14
Unclassified

**EDUCATIONAL ATTAINMENT AND LABOUR MARKET OUTCOMES
IN SOUTH AFRICA, 1994-2010**

ECONOMICS DEPARTMENT WORKING PAPERS No 1022

by Nicola Branson and Murray Leibbrandt

All Economics Department Working Papers are available through OECD's Internet website at
www.oecd.org/eco/workingpapers

JT03334881

Complete document available on OLIS in its original format

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

English - Or. English

ABSTRACT/RÉSUMÉ

Educational attainment and labour market outcomes in South Africa, 1994-2010

In this paper we document the impact of education levels on labour market outcomes from 1994 to 2010 using national household survey data. We show that higher levels of education are strongly rewarded in the labour market in terms of earnings and that a tertiary qualification improves an individual's prospects of employment. While the premium for matric and incomplete secondary has fallen marginally over the period, the premium to tertiary has risen, especially for women. Differences in the reward to education level are evident for Africans versus the overall population, between urban and rural areas and for younger versus older workers. In particular, the premium to tertiary education has increased at a higher rate for Africans than for the overall population.

This Working Paper relates to the *2013 OECD Economic Survey of South Africa*, <http://www.oecd.org/eco/surveys/listofeconomicsurveysofsouthafrica.htm>.

JEL codes: I24, J21, J31

Keywords: earnings; employment; education; South Africa; national household survey data

+++++

L'éducation et son rendement sur le marché du travail en Afrique du Sud, 1994-2010

Dans cette étude nous examinons les rendements de l'éducation sur le marché du travail entre 1994 et 2010 à l'aide d'enquêtes-ménages nationales. Nous montrons que le rendement de l'éducation supérieure en termes de salaire est très élevé et qu'un diplôme du supérieur augmente également la probabilité d'être employé. Alors qu'un niveau d'éducation secondaire a eu des rendements en légère baisse au cours de la période, le rendement de l'éducation tertiaire a augmenté, surtout pour les femmes. A niveau éducatif donné, des différences de rendement sont observées entre les Africains et la population totale, les zones rurales et urbaines et entre les jeunes et les séniors. En particulier, les rendements de l'éducation tertiaire a augmenté plus rapidement pour les Africains que pour la population totale.

Ce Document de travail se rapporte à l'*Étude économique de l'OCDE de l'Afrique du Sud 2013*, <http://www.oecd.org/fr/eco/etudeseconomiquesparpays/listofeconomicsurveysofsouthafrica.htm>.

Codification JEL: I24 ; J21 ; J31

Mots clés: salaire ; emploi ; éducation ; Afrique du Sud ; enquête ménage nationale

© OECD (2013)

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for commercial use and translation rights should be submitted to rights@oecd.org.

TABLE OF CONTENTS

1. Introduction	5
2. Private returns to education.....	7
2.1. Data and sample characteristics.....	7
2.2. Model specifications.....	14
2.3. Regression results of the private returns to education	15
2.3.3. Testing for presence of sample selection bias.....	21
3. Concluding remarks	22
Bibliography	24
<i>Appendix 1. Survey sample information and variable definitions</i>	<i>25</i>
Survey sample information:.....	25
Labour force variables:.....	25
Education variables:.....	26
<i>Appendix 2. Other figures</i>	<i>28</i>
Tables	
1. Sample summary characteristics by year and gender.....	8
2. Mean sample characteristics by education category – males – 1994 2002 2010	10
3. Mean sample characteristics by education category –females – 1994 2002 2010	11
Figures	
1. Real earnings for males and females – all earners, all earners with outliers trimmed and wage earners with outliers trimmed	12
2. Ratio of male monthly earnings to female monthly earnings by survey year	13
3. Relative earnings for incomplete secondary, matric and tertiary compared to primary - basic and controls model– all years.....	15
4. Increased probability of employment for incomplete secondary, matric and tertiary relative to primary– basic and controls model – all years	17
5. Relative earnings for incomplete secondary, matric and tertiary compared to primary – African only versus national population– all years.....	19
6. Relative earnings for incomplete secondary, matric and tertiary compared to primary – Urban and rural subsamples.....	20
7. Relative earnings for incomplete secondary, matric and tertiary compared to primary – younger (15-30) and older workers (31-64)	21
8. Comparison of education level coefficients from Heckman selection model and control model	22
A1. Classification of highest education level into years of education.....	26
A2. Number of observations classified as outliers (real earnings in 1994 rands > 56 475) by survey year	27
B1. Relative earnings for matric and tertiary compared to incomplete secondary or less – African only versus national population – females only, all years	28
B2. Relative earnings for incomplete secondary, matric and tertiary compared to primary – Urban and rural subsamples.....	29
B3. Relative earnings for incomplete secondary, matric and tertiary compared to primary – younger (15-30) and older workers	30

EDUCATIONAL ATTAINMENT AND LABOUR MARKET OUTCOMES IN SOUTH AFRICA , 1994-2010

By Nicola Branson and Murray Leibbrandt¹

1. Introduction

Reducing poverty and inequality are key challenges in South Africa. Access to the labour market and earnings differentials are the main drivers of poverty and inequality in South Africa (Leibbrandt *et al.*, 2010). Educational attainment and quality of education play an important role in determining labour market success but remain strongly correlated with socioeconomic status (Spaull, 2010). Understanding the relationship between education and the labour market and how this relationship has changed over time therefore forms an important part of the discussion.

The end of apartheid in 1994 brought about many changes in the South African labour market. Labour force participation increased significantly² in the late 1990s and early 2000s and employment, though growing, could not keep pace (Branson and Wittenberg, 2007). The increase in participation was primarily driven by an increase in African participation, especially among females, the youth and those with less education. Kingdon and Knight (2008) and Casale and Posel (2002) attribute the increase in female participation to the decrease in marriage rates, increase in single person households, improvements in educational attainment and the loss of male employment. The rise in young participants can also, in part, be attributed to the Department of Education discussion of a policy in 1995 to reduce the number of over-age learners in schools. While the plan was for learners two or more years older than the appropriate age for their grade to attend Further Education and Training colleges (FETs), evidence showed that the proposed policy primarily resulted in an increase of low-skilled workers (most with incomplete secondary education) entering the labour market, with no increase in FET enrolment (Burger and von Fintel, 2009).

-
1. Nicola Branson is a postdoctoral research fellow in the Southern Africa Labour and Development Research Unit (SALDRU) at the University of Cape Town. Murray Leibbrandt is Professor of Economics and Director of SALDRU at the University of Cape Town and the National Research Foundation Research Chair of Poverty and Inequality. Emails: nicola.branson@gmail.com, murray.leibbrandt@uct.ac.za. The authors are grateful to Geoff Barnard and Fabrice Murtin for helpful comments and suggestions. This paper was commissioned by the OECD Economics Department in the context of the *2013 Economic Survey of South Africa*. The views expressed in this study are those of the authors, and are not necessarily those of the OECD or of its member countries
 2. Burger and Yu (2006) and Kerr and Wittenberg (2012) show that much of the increase in the labour force in the late 1990s is a function of changes in the sample design and questionnaire across the surveys. In particular, later survey instruments were better designed to capture self-employment and the informal sector (Burger and Yu, 2006) and the sampling frame adjusted to include secondary households at sampling points (Kerr and Wittenberg, 2012).

The formulation and implementation of new legislations and policies played a part by changing the incentives and vulnerabilities of certain sectors of the labour market. The Basic Conditions of Employment Act of 1997 and the Employment Conditions Commission instituted basic employment conditions and minimum wages to protect some vulnerable sectors. Sectors covered by the commission include domestic work, agriculture, forestry, contract cleaning, private security, civil engineering, hospitality and tourism, the taxi sector and wholesale and retail trade (Leibbrandt *et al.*, 2010). Many of the people affected by this legislation are those in elementary employment, often female, and hence these changes in policy provide another explanation for the increase in participation seen in the 1990s-2000s. Strict labour legislation has also led to a preference for contract and casual workers especially among the most vulnerable (Leibbrandt *et al.*, 2010). Casual employment doubled between 1993 and 2008 for Africans while it grew 22% for whites over the same period (*ibid*).

This together with the natural shift away from low-skilled work such as agriculture and the increase in demand for skills to support the financial and other service industries, means highly skilled workers are in demand and of short supply.

Inequalities in the education system continue to feed divisions in labour market success by socioeconomic class. There has been a rapid increase in educational attainment in the past three decades, yet much of the increase is at the secondary, and often, incomplete secondary education level. Only marginal increases in the share of the population with tertiary education are evident in the data and the pool of matriculants furthering their education has in fact decreased. In addition, higher educational achievement remains starkly delineated by socioeconomic class. These shifts in educational attainment reflect deep inequalities in access to quality education. While 25% of the population attend schools of acceptable standard and achieve at the appropriate level, the majority of the population receive very poor quality education and achieve well below the expected level on both national and international tests (Spaull, 2012).

Aggregate wages, once the data are made consistent, have been fairly stable over the period. The raw data shows a decline in real wages between 1995 and 2005, however Burger and Yu (2006) show that this apparent decline in real wages is in fact a function of changes in the sample design and questionnaire across the surveys. In particular, they note that the decline is a function of better capture of self-employment and informal sector earnings and that once these groups are excluded, aggregate real earnings are stable between 1995 and 2005. They also point out that the presence of outliers in some surveys are influential and recommend trimming outliers over ZAR 85 000 per month in 2000 rand terms.

Racial wage discrimination remains a characteristic of the South African labour market (Rospabe, 2002; Chamberlain and van der Berg, 2002; Burger and Jafta, 2006; Burger and van der Berg, 2011). However Rospabe (2002) finds a decline in white-African employment discrimination between 1993 and 1999, but a slight increase in wage and occupation discrimination. Burger and Jafta (2006) similarly find no decline in the unexplained part of the racial wage gap between 1995 and 2005. They note that Affirmative Action appears to have been ineffective except for a small group of top earners (Burger and Jafta, 2006). Chamberlain and van der Berg (2002) and Burger and van der Berg (2011) show that differences in educational quality explain a large portion of the racial wage gap observed between population groups in South Africa. These authors attribute close to 60% of the African-white wage gap to differences in education quality and find that up to 80% of the gap can be explained by observable characteristics. Each study concludes that improvements in the quality of schooling provided to the majority of the population would be the most effective means to reduce the racial gap.

Together these factors have adjusted the composition of the South African labour market over time. In this paper we investigate what this means for private employment and earnings premiums to education level. We use seventeen years of national household survey data to document the relationship between

education level and employment and earnings between 1994 and 2010. Using a simple but consistent specification of this relationship enables us to assess trends and theorise over possible reasons for these changes. This paper therefore lays the foundation for more in-depth work, especially regarding choice of specification and the inclusion of measures of school quality, on the returns to education.

We show that higher levels of education are strongly rewarded in the labour market in terms of earnings and that, in addition, a tertiary qualification improves an individual's prospects of employment. We find that while the premium for matric and incomplete secondary has fallen marginally over the period investigated, the premium to tertiary has risen, especially for women. We illustrate differences in reward to education level for Africans versus the overall population, between urban and rural areas and younger versus older workers. The premium to tertiary is increasing at a higher rate for Africans than the overall population increase. This signals that Africans with tertiary education are, in particular, increasingly valued in the labour market. A similar pattern is apparent between younger versus older workers, with premiums to tertiary increasing more strongly for older workers. This highlights the increasing vulnerability of younger cohorts of workers and shows that they have not benefited from the increase in the premium to tertiary.

The data and sample characteristics used in the analysis of the private returns to education are presented in Section 2.1, followed by the estimation specification in Section 2.2. The regression results of the private returns to education are presented in Section 2.3 with a discussion of these results in Section 2.4. Section 3 reviews the overall story.

2. Private returns to education

In this section we investigate the benefit of educational attainment for individual labour market outcomes. As such we focus exclusively on the private returns to education and no attempt is made to uncover the social returns to education.

2.1. Data and sample characteristics

We use data from seventeen national cross sectional household surveys spanning 1994 to 2010. The data is compiled from the October Household Surveys (OHS's) in the 1990s, the Labour Force Surveys (LFSs) between 2000 and 2007 and the General Household Surveys (GHSs) for 2008 through 2010. These are currently all the publicly available national household surveys that contain individual level earnings information in addition to individual and household characteristics. We use the Post Apartheid Labour Market Series (PALMS) version of the 1994-2007 data as our starting point and supplement these with GHS data in 2008, 2009 and 2010³. Most importantly for this analysis, these surveys provide individual level information on highest level of education completed. From this we can distinguish between individuals who have primary or less education, have completed some secondary education, have completed secondary education (matric) or some post-schooling studies. The sample is restricted to individuals between 15-64 in each year.

Education is compulsory in South Africa from age seven (grade 1) to age 15 or the completion of grade 9 and enrolment in these grades is almost universal. Grades 10 through 12 are referred to as further education and training (FET) since learners can choose between a vocational training route or to continue their education in the basic education system. Those who choose the vocational route complete this with a National Certificate Vocational (NCV). The nationally administered National Senior Certificate (NSC) taken in grade 12 represents the completion of basic education and continues to be the preferred choice. Enrolment rates are high, but levels of grade repetition are also high in most grades and the majority of

3. See Appendix A for survey data information and variable definitions.

learners (58%⁴) still leave the schooling system without completing a national leaving certificate (NCV or NSC) commonly known as matric. This reflects the poor quality of schooling in most South African schools such that although educational attainment has increased over the past decades, completion of matric continues to represent a large and important hurdle to cross.

Table 1. Sample summary characteristics by year and gender

	Male					Female				
	OHS 1994	OHS 1998	LFS 2002	LFS 2006	GHS 2010	OHS 1994	OHS 1998	LFS 2002	LFS 2006	GHS 2010
Age	32.54	32.60	32.71	32.87	33.17	33.03	33.13	33.39	33.79	34.31
African	0.74	0.75	0.76	0.77	0.78	0.74	0.76	0.77	0.78	0.78
Coloured	0.10	0.09	0.09	0.09	0.10	0.10	0.09	0.09	0.09	0.10
Indian	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
White	0.14	0.12	0.11	0.11	0.10	0.13	0.12	0.11	0.10	0.10
Married	0.45	0.44	0.41	0.39	0.40	0.48	0.45	0.41	0.39	0.41
Years of education	8.19	8.29	8.74	9.15	9.55	7.97	8.11	8.60	8.99	9.46
Primary	0.28	0.27	0.26	0.21	0.18	0.28	0.26	0.24	0.20	0.17
Incomplete Secondary	0.41	0.39	0.40	0.42	0.44	0.42	0.40	0.41	0.42	0.43
Matric	0.16	0.18	0.21	0.23	0.25	0.15	0.17	0.19	0.23	0.25
Tertiary	0.07	0.07	0.08	0.09	0.10	0.06	0.06	0.08	0.08	0.10
Informal housing	0.24	0.25	0.25	N/A	N/A	0.26	0.27	0.25	N/A	N/A
Number of children in household	2.04	1.93	1.80	1.60	1.52	2.53	2.48	2.31	2.16	2.05
Number of working adults in household	1.29	1.17	1.16	1.20	1.23	1.19	1.07	1.07	1.12	1.12
Number of pension aged adults in household	0.26	0.27	0.25	0.23	0.24	0.30	0.30	0.29	0.27	0.27
Urban	0.56	0.62	0.62	N/A	N/A	0.53	0.58	0.59	N/A	N/A
Western Cape	0.11	0.11	0.11	0.12	0.12	0.10	0.10	0.11	0.11	0.11
Eastern Cape	0.13	0.12	0.12	0.12	0.13	0.15	0.15	0.14	0.13	0.13
Northern Cape	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Free State	0.07	0.07	0.06	0.06	0.06	0.07	0.06	0.06	0.06	0.06
KwaZulu-Natal	0.20	0.20	0.20	0.20	0.19	0.22	0.22	0.21	0.21	0.20
North West	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Gauteng	0.22	0.24	0.24	0.24	0.23	0.19	0.19	0.21	0.21	0.21
Mpumalanga	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
Limpopo	0.09	0.09	0.09	0.10	0.11	0.11	0.11	0.11	0.12	0.12
Real monthly earnings (1994 rands)	2753.75	2546.05	2120.50	2343.58	2604.45	1584.78	1647.70	1432.11	1683.20	2059.70
Real monthly earnings - wage earners only	1995.68	1994.89	1977.51	2223.11	2504.02	1733.29	1471.11	1517.16	1747.42	2122.91
Monthly earnings	2753.75	3461.69	3510.09	4547.45	6737.73	1584.78	2240.26	2372.53	3265.60	5328.46
Economically Active	0.69	0.68	0.74	0.74	0.78	0.51	0.52	0.65	0.65	0.70
Unemployed	0.17	0.20	0.25	0.23	0.24	0.21	0.23	0.31	0.30	0.33
Employed	0.51	0.48	0.49	0.51	0.53	0.30	0.28	0.34	0.36	0.37
Job Industry:										
Agriculture, hunting, forestry and fishing	0.19	0.12	0.14	0.10		0.08	0.07	0.11	0.09	
Mining and quarrying	0.05	0.09	0.08	0.06		0.00	0.00	0.00	0.00	
Manufacturing	0.19	0.16	0.17	0.16		0.14	0.13	0.12	0.11	
Utilities	0.01	0.02	0.01	0.01		0.00	0.01	0.00	0.00	
Construction	0.07	0.09	0.08	0.12		0.01	0.01	0.01	0.02	
Trade	0.16	0.17	0.18	0.22		0.20	0.22	0.23	0.27	
Transport	0.08	0.08	0.07	0.07		0.02	0.03	0.02	0.02	
Finance	0.05	0.09	0.10	0.10		0.07	0.08	0.09	0.10	
Services	0.21	0.15	0.14	0.14		0.48	0.27	0.23	0.23	
Domestic Services	0.00	0.02	0.03	0.03		0.00	0.18	0.18	0.16	

Note: The table presents the weighted mean sample characteristics by survey year for males and females separately. The sample is restricted to 15-64 year olds. OHS refers to October Household Survey, LFS to Labour Force Survey and GHS to General Household Survey. Means weighted using the cross entropy weight. N/A represents not available.

Table 1 presents a summary of the characteristics of the sample separately for males and females for five of the survey years. Although levels of educational attainment are very similar between males and females, their labour market choices and experiences differ substantially. We therefore report separate analyses for males and females throughout this paper. Tables 2 and 3 further disaggregate the summary characteristics by education level for males and females respectively. Average educational attainment was 9.5 years for both males and females in 2010, indicating that the majority of South Africans complete more than the compulsory level of schooling but that completion of school (grade 12) is still low. Important to this analysis is the improvement in educational attainment over the period. The increase in the proportion with matric is most substantial. In 1994, only 15% of the female and 16% of the male sample had completed matric, by 2010 25% of both the female and male sample had completed matric. There has been

4. Only 42% of 25-29 year olds in 2008 had completed grade 12 (Branson and Lam, 2010).

a small increase in the proportion of males (females) with completed tertiary education from 7% (6%) in 1994 to 10% (10%) in 2010, however the fraction of matriculants continuing on to tertiary has declined.

Further examination of Table 1 reflects the increasingly difficult labour market conditions South Africans face. The large increase in labour force participation⁵ between the late 1990s and early 2000s, especially among females, is evident in the table. Bearing in mind the concerns raised around the data (in the introduction), we see an increase in economic activity from 68% to 74% between 1998 and 2002 for males and from 52% to 65% for females. Employment has not kept pace with this increase in labour force participation. The proportion of males employed increased only slightly from 51% to 53% over the period and the proportion of females employed increased by seven percent from 30% in 1994 to 37% in 2010. This was not enough to absorb the increases in labour force participation and unemployment is shown to increase from 17% for males and 21% for females in 1994 to 24% for males and 33% for females in 2010.

Tables 2 and 3 show that economic activity increased within each education category for females and in all but the primary category for males. The rates of increase were however much higher for females and particularly those with incomplete secondary education. While employment increased on aggregate, we see that females with incomplete secondary or less are only marginally more likely to be employed, and that males with incomplete secondary or less are actually less likely to be employed. Thus the increase in the supply of educated workers and the increase in jobs available by skill level are not in alignment.

The skill-biased growth of employment in South Africa is evident when the distribution of employment across sectors is examined, especially by education level. In Table 1 we see an overall increase in the share of workers employed in trade, construction and finance, a decline in the share in agriculture and manufacturing, and the share of employment in the other sectors has remained fairly stable. Moving to Tables 2 and 3, we see that the majority of male workers with primary education work in agriculture and the majority of female workers with primary are in domestic services and agriculture. For incomplete secondary the main sectors of employment are trade, services and manufacturing, for matric the main sectors are trade, finance and services and for tertiary the predominant sector of employment is services with finance second. Thus those with primary and incomplete secondary education work primarily in sectors that are contracting while those with matric and tertiary are predominantly in the growing sectors of the economy.

5. Broad definition of labour force participation used.

Table 2. Mean sample characteristics by education category – males – 1994 2002 2010

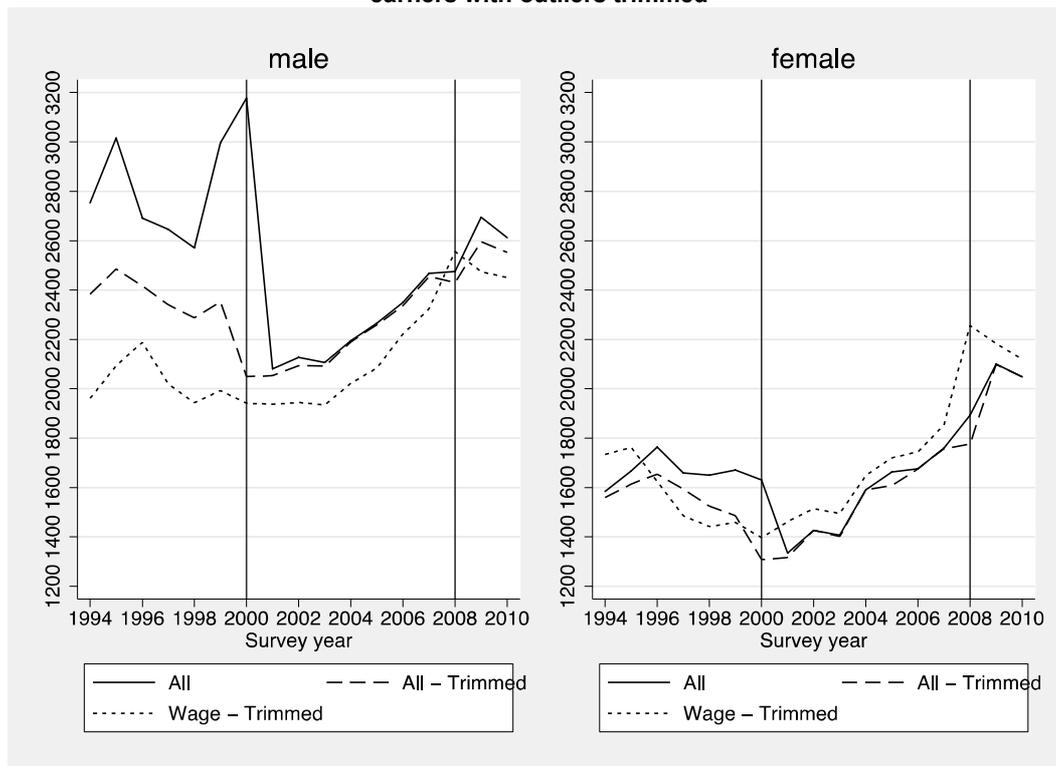
	OHS 1994				LFS 2002				GHS 2010			
	Incomplete				Incomplete				Incomplete			
	Primary	Secondary	Matric	Tertiary	Primary	Secondary	Matric	Tertiary	Primary	Secondary	Matric	Tertiary
Age	34.89	30.15	31.37	36.91	36.27	29.29	31.53	37.80	38.25	30.07	32.25	38.34
African	0.89	0.74	0.53	0.39	0.90	0.79	0.62	0.43	0.91	0.83	0.68	0.49
Coloured	0.10	0.11	0.08	0.05	0.09	0.10	0.09	0.06	0.08	0.10	0.10	0.08
Indian	0.01	0.04	0.07	0.04	0.01	0.03	0.06	0.04	0.01	0.02	0.05	0.06
White	0.01	0.11	0.33	0.53	0.01	0.08	0.23	0.46	0.01	0.05	0.17	0.37
Married	0.48	0.38	0.46	0.71	0.46	0.31	0.41	0.66	0.45	0.32	0.41	0.65
Years of education	3.93	9.45	12.00	14.17	4.31	9.48	12.00	14.09	4.44	9.63	12.00	14.00
Urban	0.38	0.61	0.76	0.81	0.42	0.63	0.80	0.86				
Western Cape	0.09	0.11	0.13	0.15	0.09	0.11	0.12	0.17	0.09	0.11	0.13	0.17
Eastern Cape	0.18	0.12	0.08	0.08	0.17	0.11	0.08	0.09	0.19	0.13	0.10	0.09
Northern Cape	0.03	0.02	0.01	0.01	0.03	0.02	0.02	0.02	0.03	0.02	0.01	0.01
Free State	0.09	0.07	0.05	0.05	0.07	0.06	0.06	0.06	0.07	0.05	0.06	0.05
KwaZulu-Natal	0.18	0.21	0.23	0.20	0.20	0.20	0.20	0.16	0.20	0.20	0.21	0.14
North West	0.11	0.07	0.06	0.05	0.10	0.08	0.07	0.05	0.11	0.08	0.07	0.05
Gauteng	0.15	0.23	0.28	0.32	0.15	0.24	0.33	0.34	0.13	0.21	0.30	0.34
Mpumalanga	0.09	0.06	0.07	0.05	0.08	0.06	0.06	0.05	0.07	0.07	0.06	0.06
Limpopo	0.08	0.11	0.09	0.08	0.11	0.10	0.06	0.08	0.12	0.13	0.06	0.09
Informal housing	0.37	0.21	0.09	0.04	0.38	0.25	0.13	0.03				
Number of children in household	2.33	2.11	1.52	1.31	2.18	1.90	1.32	1.05	1.96	1.67	1.11	0.90
Number of working adults in household	1.09	1.26	1.61	1.75	1.00	1.10	1.36	1.57	0.91	1.13	1.47	1.70
Number of pension aged adults in household	0.29	0.27	0.23	0.17	0.30	0.25	0.22	0.18	0.29	0.24	0.22	0.16
Real monthly earnings (1994 rands)												
All	957.26	2238.05	4111.80	7343.40	740.97	1397.31	2574.14	6303.44	847.25	1345.31	2627.94	7487.34
Trimmed	906.76	2027.06	3661.20	5760.33	740.97	1397.31	2574.14	6059.17	847.25	1345.31	2627.94	7130.16
Excl self-employed	798.08	1735.26	2943.92	4778.46	737.45	1350.54	2345.83	5639.83	872.81	1322.79	2465.67	6891.65
Monthly earnings	957.26	2238.05	4111.80	7343.40	1224.51	2313.48	4255.25	10448.97	2191.85	3480.32	6798.50	19369.79
Economically Active	0.71	0.58	0.79	0.90	0.72	0.65	0.87	0.93	0.69	0.70	0.91	0.96
Unemployed	0.22	0.15	0.16	0.04	0.26	0.26	0.29	0.10	0.27	0.27	0.25	0.07
Employed	0.48	0.43	0.64	0.86	0.46	0.38	0.58	0.84	0.42	0.43	0.66	0.89
Union	0.22	0.41	0.38	0.33	0.30	0.36	0.38	0.46				
Job Industry:												
Agriculture, hunting, forestry and fishing	0.43	0.10	0.03	0.03	0.30	0.10	0.05	0.03				
Mining and quarrying	0.05	0.05	0.05	0.03	0.12	0.09	0.05	0.03				
Manufacturing	0.13	0.25	0.20	0.14	0.11	0.20	0.21	0.14				
Utilities	0.01	0.02	0.02	0.02	0.01	0.01	0.01	0.01				
Construction	0.09	0.08	0.04	0.04	0.11	0.10	0.05	0.02				
Trade	0.11	0.21	0.20	0.09	0.14	0.21	0.23	0.10				
Transport	0.05	0.11	0.09	0.04	0.05	0.09	0.09	0.05				
Finance	0.01	0.03	0.10	0.15	0.02	0.08	0.15	0.22				
Services	0.12	0.16	0.27	0.47	0.07	0.09	0.16	0.39				
Domestic Services	0.00	0.00	0.00	0.00	0.07	0.03	0.01	0.00				

Table 3. Mean sample characteristics by education category – females – 1994 2002 2010

	OHS 1994				LFS 2002				GHS 2010			
	Incomplete				Incomplete				Incomplete			
	Primary	Secondary	Matric	Tertiary	Primary	Secondary	Matric	Tertiary	Primary	Secondary	Matric	Tertiary
Age	37.13	29.74	30.85	35.62	39.25	29.46	30.64	36.64	43.12	30.57	31.81	37.40
African	0.88	0.75	0.51	0.43	0.89	0.80	0.61	0.53	0.91	0.84	0.66	0.55
Coloured	0.10	0.11	0.08	0.04	0.09	0.10	0.09	0.06	0.08	0.10	0.11	0.08
Indian	0.02	0.03	0.05	0.03	0.01	0.03	0.05	0.03	0.01	0.02	0.04	0.05
White	0.00	0.11	0.36	0.49	0.01	0.08	0.25	0.37	0.01	0.04	0.19	0.32
Married	0.55	0.40	0.50	0.66	0.47	0.33	0.42	0.59	0.46	0.34	0.43	0.58
Years of education	3.84	9.47	12.00	13.87	4.02	9.52	12.00	13.84	4.09	9.69	12.00	13.94
Urban	0.34	0.59	0.75	0.76	0.39	0.61	0.77	0.82				
Western Cape	0.08	0.10	0.13	0.14	0.08	0.11	0.13	0.15	0.07	0.11	0.13	0.15
Eastern Cape	0.19	0.14	0.09	0.11	0.17	0.14	0.09	0.10	0.16	0.14	0.10	0.11
Northern Cape	0.03	0.02	0.01	0.02	0.03	0.02	0.02	0.01	0.03	0.02	0.01	0.01
Free State	0.08	0.06	0.05	0.05	0.07	0.06	0.06	0.06	0.07	0.06	0.05	0.05
KwaZulu-Natal	0.21	0.21	0.24	0.20	0.23	0.21	0.20	0.16	0.22	0.20	0.21	0.16
North West	0.09	0.07	0.06	0.05	0.09	0.07	0.07	0.05	0.10	0.08	0.06	0.05
Gauteng	0.11	0.22	0.26	0.31	0.11	0.21	0.30	0.31	0.11	0.19	0.30	0.32
Mpumalanga	0.09	0.06	0.07	0.04	0.08	0.06	0.06	0.05	0.09	0.07	0.06	0.06
Limpopo	0.13	0.11	0.08	0.09	0.13	0.12	0.08	0.10	0.15	0.14	0.07	0.10
Informal housing	0.38	0.23	0.09	0.05	0.37	0.25	0.12	0.03				
Number of children in household	2.91	2.57	1.81	1.59	2.72	2.45	1.76	1.34	2.45	2.32	1.59	1.15
Number of working adults in household	0.96	1.18	1.55	1.83	0.92	1.00	1.29	1.50	0.78	1.00	1.41	1.62
Number of pension aged adults in household	0.33	0.29	0.25	0.18	0.34	0.29	0.26	0.22	0.35	0.26	0.26	0.19
Real monthly earnings (1994 rands)												
All	583.65	1271.40	2303.64	3415.00	372.01	933.16	1869.05	3690.51	541.09	893.41	2157.20	4881.44
Trimmed	583.65	1222.69	2303.64	3348.46	372.01	933.16	1869.05	3690.51	541.09	893.41	2157.20	4881.44
Excl self employed	672.88	1317.07	2150.58	3169.72	389.94	996.72	1899.56	3624.74	549.77	964.59	2151.03	4812.77
Monthly earnings	583.65	1271.40	2303.64	3415.00	614.77	1542.18	3096.73	6123.52	1399.79	2311.26	5580.69	12628.31
Economically Active	0.49	0.44	0.65	0.80	0.61	0.58	0.79	0.87	0.57	0.64	0.85	0.92
Unemployed	0.26	0.20	0.21	0.05	0.29	0.33	0.38	0.16	0.31	0.38	0.35	0.13
Employed	0.23	0.24	0.45	0.75	0.32	0.25	0.40	0.71	0.26	0.26	0.50	0.79
Union	0.15	0.30	0.29	0.30	0.11	0.25	0.30	0.53				
Job Industry:												
Agriculture, hunting, forestry and fishing	0.21	0.05	0.01	0.01	0.25	0.10	0.03	0.01				
Mining and quarrying	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01				
Manufacturing	0.12	0.20	0.12	0.05	0.09	0.16	0.14	0.05				
Utilities	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.01				
Construction	0.01	0.01	0.02	0.01	0.02	0.01	0.01	0.00				
Trade	0.15	0.27	0.23	0.08	0.20	0.29	0.29	0.09				
Transport	0.00	0.02	0.03	0.01	0.00	0.02	0.05	0.03				
Finance	0.01	0.04	0.16	0.10	0.01	0.06	0.17	0.15				
Services	0.50	0.40	0.40	0.73	0.06	0.16	0.25	0.65				
Domestic Services	0.00	0.00	0.00	0.00	0.37	0.19	0.05	0.00				

Notes to Table 2 and 3: The table presents the weighted mean sample characteristics by survey year and education category for males and females respectively. The sample is restricted to 15-64 year olds. OHS refers to *October Household Survey*, LFS to *Labour Force Survey* and GHS to *General Household Survey*. Means weighted using the cross entropy weight.

Figure 1. Real earnings for males and females – all earners, all earners with outliers trimmed and wage earners with outliers trimmed



Note: The figure presents mean real earnings in 1994 rands over time separately for males and females. All includes both wage and self-employment earnings. Wage includes only those with wage employment. Trimmed indicates that values over ZAR 56 475 were excluded as outliers (as per Burger and Yu, 2006). Means weighted using the cross entropy weights. 'Wage-trimmed' excludes those with self-employment and hence its average can be larger than the average for 'All'.

Figure 1 presents real earnings for all earners, all earners with outliers trimmed and wage earners only with outliers trimmed, separately for males and females for each survey year. We use the Burger and Yu (2006) cut-off converted to 1994 rand terms to trim outliers in each year⁶. Similar to Burger and Yu (2006), once self-employed earnings are excluded and outliers trimmed, real earnings of both males and females decrease between 1994 and 2000 but thereafter increase. The horizontal lines on the figure represent shifts in the survey source. Interestingly, the exclusion of self-employment earnings decreases mean earnings in most years for men and increases earnings in most years for females likely reflecting differences in self-employment occupation type and education level between males and females⁷.

We would expect trends in earnings to differ by education level since the demand and supply of certain skills is not constant across education level. Tables 2 and 3 show that this is indeed true. In Table 2, Males with incomplete or completed secondary experienced stable if not declining real wages between 2002 and 2010⁸, while the real earnings of those with primary or tertiary increased over the same period. Table 3 presents similar statistics for females. Females with tertiary experienced a substantial (32%) increase in real earnings between 2002 and 2010. The earnings of females with primary education also

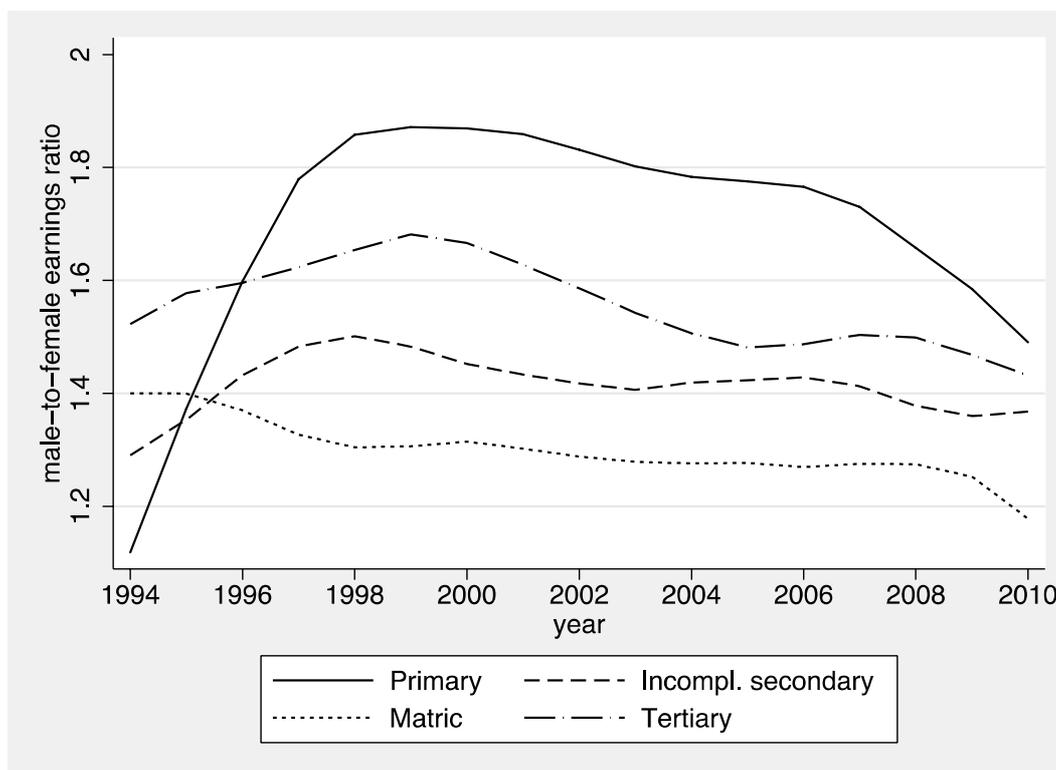
6. Appendix A presents a table of the number of observations trimmed in survey year.

7. The vast majority of females in self-employment have primary or no education and are in elementary occupations while males in self-employment are more evenly distributed across education categories and occupation types.

8. We discuss trends between 2002 and 2010 to avoid concerns of inconsistency between the OHS and LFS data sets. See footnote 2.

increased substantially (40%). While in Tables 2 and 3 we see that both males and females with incomplete secondary experienced a small decline in real wages, the earnings of females with matric increased by about 15% compared to an increase of only about 5% for males with matric. As a result the gap between those with tertiary and those with secondary education has increased and the gender gap has decreased.

Figure 2. Ratio of male monthly earnings to female monthly earnings by survey year



Note: The figure presents the ratio of male monthly earnings to female monthly earnings over time for each education category. Estimates weighted using the cross entropy weights. The strange estimates prior to 2000 for the primary group reflect changes in survey methodology. See Footnote 1.

Figure 2 investigates the difference in male versus female earnings further. The ratio of male-to-female wage earnings is displayed for each education category and survey year. Ignoring the first three survey years as these trends are likely to reflect measurement changes in the surveys, we see that there is a systematic pattern across the education categories; the gender gap is largest in each year for the primary group, followed by tertiary, incomplete secondary and matric. The gender gap is shown to decrease within each education category between 1998 and 2010, but the decrease is larger for the groups with large gaps.

Tables 2 and 3 provide further details of differences in labour market and individual characteristics of the population by education level. The demographic, racial and geographic characteristics of individuals vary by education level. Each subsequent level of education has a higher representation of whites and lower representation of Africans. For example in 1994, 89% of those in the primary group were African and 1% were white, compared to 39% African and 53% white in the tertiary group. There have been some improvements in African representation in the tertiary group over the period. In 2010, 49% of males and 55% of females with tertiary were African. Those with higher levels of education are also more likely to live in urban areas. This difference is most distinct for the primary group relative to the other groups.

In the next section we specify the model used to calculate the private returns to education level for each year 1994 through 2010.

2.2. Model specifications

The initial model presented is a basic semi log linear wage regression with a quadratic in age (in single years) and education level dummies. Age and age squared are included to account for age and experience since the high rates of grade repetition that characterise the South African schooling system and the high rates of unemployment, especially youth unemployment, limits the appropriateness of the traditional experience specification in the South African context.

$$\ln(Wage)_i = \alpha_i + \beta_1 Age_i + \beta_2 Age_i^2 + \gamma_1 IncSec_i + \gamma_2 Matric_i + \gamma_3 Tert_i + \mu_i \quad (1)$$

where the dependent variable is the natural logarithm of monthly earnings in South African Rands. The model is run for each survey year and an indicator of the month of the survey included for years with more than one survey. This specification is extended to include controls first just for population group (with African the reference category) and then adds additional controls for marital status, number of children 0-6 in the household, number of children 7-17 in the household and number of working adults in the household marital status, number of children and an indicator for urban residence where available (1994-2004). Separate models are run for males and females.

Next we investigate the relationship between education level and employment probability. We present estimates from a linear probability model of the probability of employment given that the individual participates in the labour force.

$$Employed_i = \alpha_i + \beta_1 Age_i + \beta_2 Age_i^2 + \gamma_1 IncSec_i + \gamma_2 Matric_i + \gamma_3 Tert_i + \mu_i \quad (2)$$

where the dependent variable indicates whether individuals in the labour force have any form of employment. Here again we present the increased propensity to be in employment that incomplete secondary, matric and tertiary afford individuals over those with primary or less education graphically. The basic specification includes a quadratic in age, and the controls model supplements this with controls for coloured, Indian, white, married, divorced, widowed, number of children 0-6 in the household, number of children 7-17 in the household, number of working adults in the household and an indicator of urban residence when available.

We investigate differences in the education coefficients from equation 1 for certain groups of interest. In particular, we compare the education premium for Africans only versus the national population, between those living in urban and rural areas, and between those in a younger (15-30) subsample compared to an older subsample (31-64).

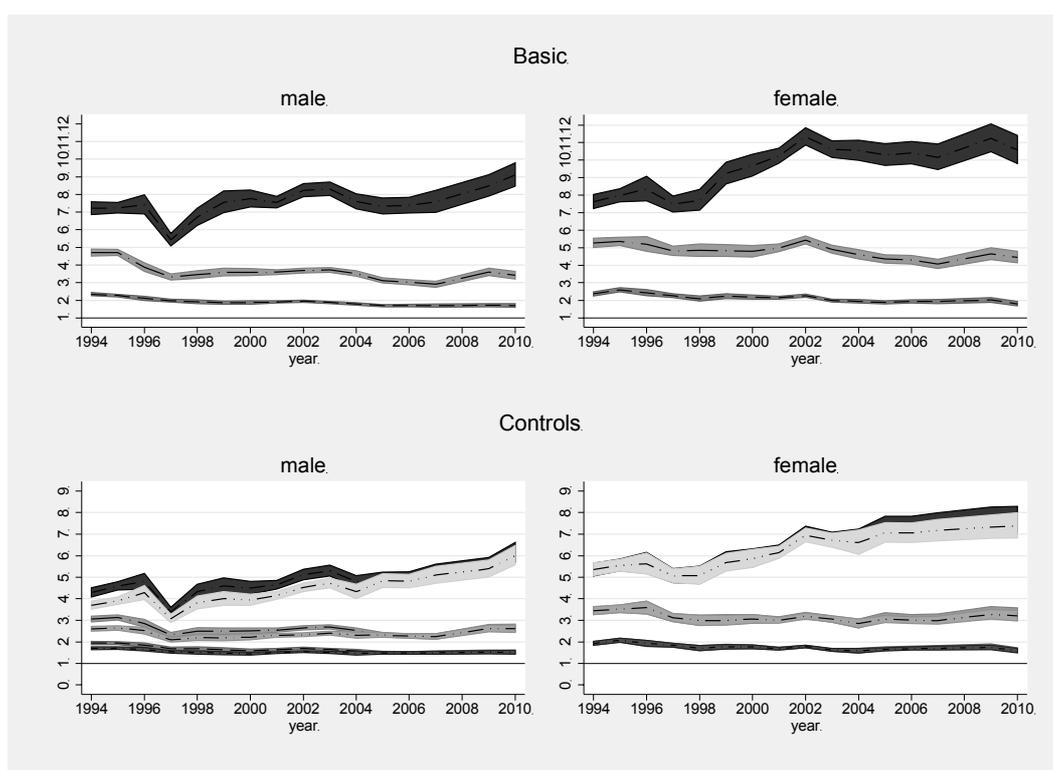
Equation 1 estimates the premium to education level for those with employment. Since people who work are not a random sample from the population, estimating premiums on a sample of individuals with employment can introduce bias. In an attempt to account for selection into employment we include estimates from a Heckman selection model estimated using maximum likelihood estimation (Wooldridge, 2002). A probit version of equation 2 with the full set of controls forms the specification for the selection equation and the version of equation 1 with population group controls and an indicator of urban residence forms the specification for the main equation. In this way we assume that marital status, number of children and number of employed adults affects the probability of employment but not earnings once employed. This is clearly unrealistic, and the estimation of equation 1 with the full set of controls shows that these characteristics do statistically influence earnings. They do not however effect the size or significant of the education level coefficients. We therefore complete the exercise for illustrative purposes.

2.3. Regression results of the private returns to education

2.3.1. Basic wage and employment regressions over time

Figure 3 presents the educational category coefficients by year for both the initial (basic) and extended specification (controls). The relative monthly earnings of those with incomplete secondary, matric and tertiary compared to with primary⁹ are presented for years 1994 through 2010 for males and females separately. The full regression results can be found in appendix B. The top panel of Figure 3 presents estimates from the model including a quadratic in age only. The controls figure presents estimates from two models overlaid. The first specification only adds coloured, Indian and white dummies to the basic specification and hence Africans with primary or less education represent the reference category. These estimates are in the background. The foreground estimates in the control figure include population group dummies, marital status, number of children (0-6 and 7-18) and number of employed adults in the household as well as an indicator of urban residence when available (1994-2004). In each panel the top line represents the premium for tertiary, the second line the premium for matric and the bottom line the premium for incomplete secondary. Point estimates are bounded by 95% confidence bands.

Figure 3. Relative earnings for incomplete secondary, matric and tertiary compared to primary - basic and controls model- all years



Note: Basic presents the converted wage regression coefficients (and confidence bands) on incomplete secondary (bottom line), matric (middle line) and tertiary (top line) relative to primary or less from a linear regression of log earnings, controlling for a quadratic in age only. The controls figures include estimated coefficients from two models overlain. The estimates in the background come from a model including only coloured, Indian and white indicators and the estimates in the foreground from a model including coloured, Indian and white indicators, marital status, number of children and number of employed adults in the household as well as an indicator of urban residence when available. Converted coefficients = $\exp(b)$ and are therefore interpreted as the relative earnings compared to those with primary. For example, in 2010 males with tertiary earned, on average, 5 times the earnings of males with primary only. All regressions weighted using the cross entropy weights.

9. These are the transformed $\exp(\gamma_i)$ regression coefficients on each of the education category dummies.

Figure 3 shows that there is an earnings premium for each education level relative to primary; none of the education confidence bands overlap the zero line. It is also noticeable that the premium to each education category is proportionately higher for females than males. This must be viewed in the context of low earnings for female primary educated individuals. In Table 3 we saw that females with primary education earned on average ZAR 540 in 2010 compared to the average of ZAR 850 for males. We also saw that the ratio of male-to-female was highest for those with primary education, meaning that male and female earnings were most distinct within this group. Hence this higher return for females reflects the low level of earnings among females with primary education and the more similar earnings of males and females within the other education categories.

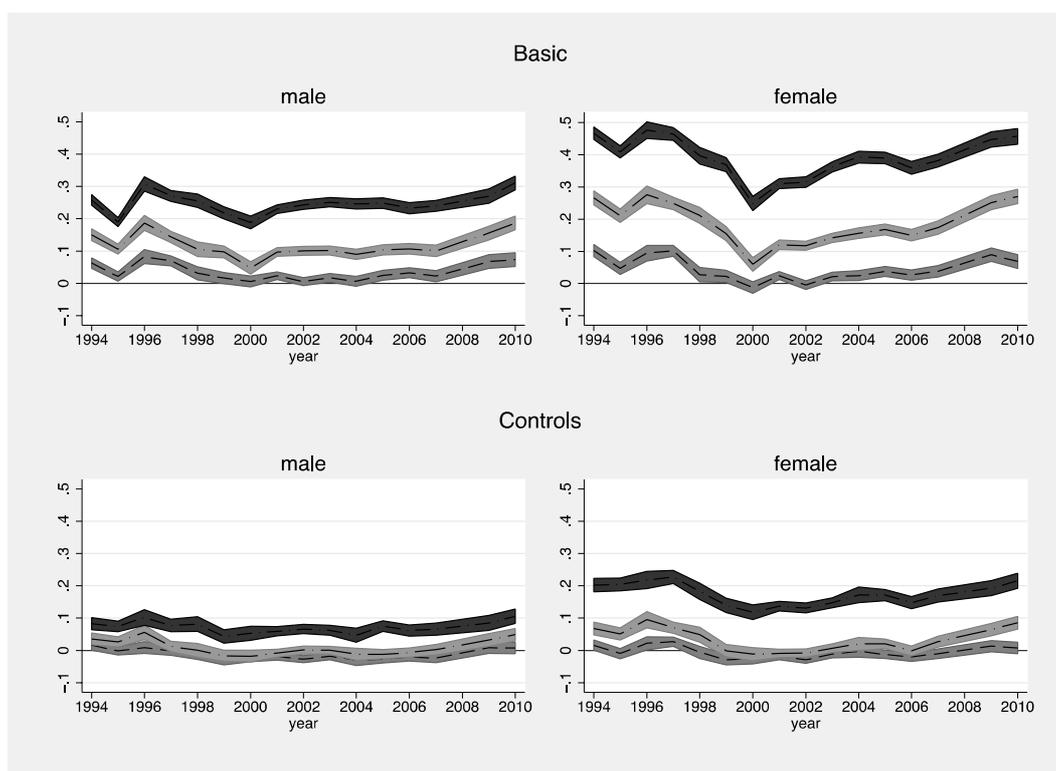
Returns to education are particularly high and strongly increasing over time for those with tertiary for both males and females. Focusing on females, we see that there was a sharp increase in the premium to tertiary between 1998 and 2002. In 1998 females with tertiary earned around eight times the earnings of females with primary or less. By 2002 this had increased to over eleven times the mean earnings of those with primary. Note, however, that data prior to 2000 comes from the OHSs and concerns have been raised about the consistency of this series over time (see footnote 2). Focusing on post-2000 data, we see that the premium to tertiary increased between 2000 and 2002, after which it declined marginally until 2007 and then increased again. Throughout this period tertiary earnings remained at over ten times the earnings of those with primary. For males the premium to tertiary in the basic model remained around 7.5 times the earnings of those with primary until 2007 (ignoring 1997), after which it increased. Premiums to matric and incomplete secondary have remained stable or declined marginally over the period for both males and females, but they remain statistically different from one in all years.

Examining the control figures in the bottom panels of Figure 3 it is clear that the racial dummies account for a large share of the education category premiums. Estimates controlling for race are just over half the size of those without race controls. Yet even once the full set of controls is included the premiums for each level of education remain positive and statistically different from zero in each year. Including controls makes the increase in the tertiary premium over time more gradual and systematic. For example, once controls for race are included the male tertiary premium increases almost year on year between 1994 and 2010 from 4 to 6 times the salary of someone with primary education only. The similarity of the overlain estimates illustrates that even though the marital status, household composition and urban indicators are significantly related to earnings they do not reduce the coefficients on the education categories in a significant way.

In this simplistic specification, the population group dummies account for more than racial discrimination. As mentioned earlier, access to quality education varies substantially by socioeconomic class and race remains a strong predictor of socioeconomic class. Other variables that are likely to affect earnings but are omitted from the specification since they are not observed in the datasets (parental and household characteristics, characteristics of immediate neighbourhoods and connections to the labour market etc) also remain strongly delineated by race in South Africa. Studies investigating the African-white earnings gap find that a large proportion of the racial gap can be attributed to observable characteristics, especially quality of education (Burger and Jafta 2006; Chamberlain and van der Berg, 2002; Burger and van der Berg, 2011; Rospabe, 2002). Thus, the reduction in the premium on each education category should not be taken as evidence of the level of racial discrimination in the labour market.

Next we investigated the relationship between education level and the odds of finding employment. Figure 4 presents the increased odds of employment from a linear probability model that incomplete secondary education, matric and tertiary afford an individual in the labour market over someone with primary or less. The top panels control only for a quadratic in age, while the bottom panels control for population group, marital status, number of children (0-6 and 7-18), number of employed adults in the household and urban residence in 1994-2004. The figure shows that the employment premium decreased between 1994 and 2000 for all education categories and increased thereafter. This is most evident for females. This will in part be explained by the increase in, or increased capture of, labour force participation at the time. Although employment did not keep pace with the increase in labour force participation we saw in Table 3 that the female employment rate improved for those with primary between 1994 and 2002, while it declined within each of the other education groups. The period also saw improved capturing of informal work that would have primarily affected low earners, predominantly with low levels of education. As such, the relative employment reward to higher levels of education decreased. It is therefore likely that the decrease between 1994 and 2000 is primarily a result of changes in the collection of the data and does not necessarily present a real trend.

Figure 4. Increased probability of employment for incomplete secondary, matric and tertiary relative to primary– basic and controls model – all years



Note: Basic presents the marginal effects (and 95% confidence bands) for incomplete secondary (bottom line), matric (middle line) and tertiary (top line) relative to primary or less from an employment probit regression controlling for a quadratic in age only. Controls, includes population group indicators, marital status, number of children under 7, between 7 and 18, number of working adults in the household and an indicator of urban residence when available. All regressions weighted using the cross entropy weights.

Once household and racial controls are taken into account, only the tertiary employment premium is statistically significant in all years. This means that only those with tertiary have an increased probability of being in employment relative to those with primary education in each year. Males with tertiary are about 10 percentage points more likely to be employed than those with primary. Females with tertiary are about 20 percentage points more likely to be employed than those with primary. The matric coefficient is small in most years and only significantly different from zero in a few years. The employment premium to matric does however appear to have increased in the last three survey years, especially for females.

2.3.2. Investigating differences across groups

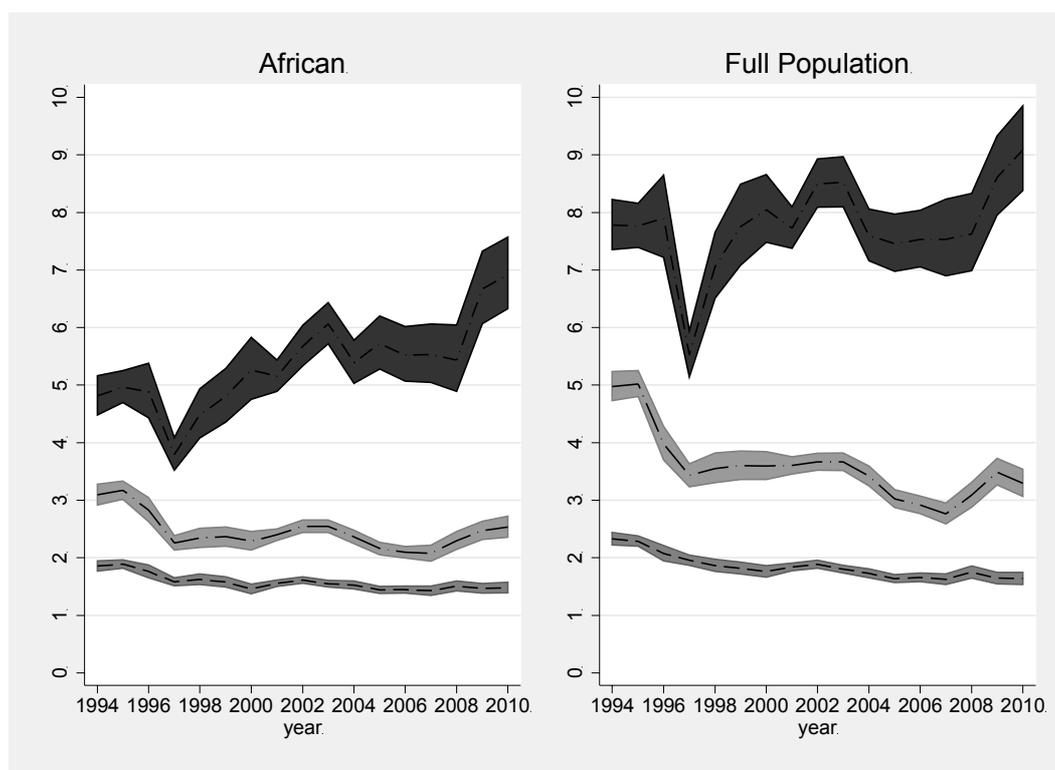
In this section we investigate differences in the education-level-earnings-premiums across different subgroups in the population. Separate regressions were run for each of the sub-samples, hence the reference category is within group and is not comparable across groups. The figures and discussion presented are for males, similar figures are supplied in Appendix B for females.

Africans versus the national population

Figure 5 presents the relative earnings for tertiary, matric and incomplete secondary compared to those with primary or less education for African males (in the left panel) and the full national population (in the right panel).

The trends in the African only figure map, as would be expected, the national trends. There are, however, some noticeable differences in the levels of the educational premiums and the rate of change over time. First, the premium for each education level is lower within the African group than the overall national population. This is particularly true for the tertiary premium. In 1994 Africans with tertiary earn about five times the salary of Africans with primary. The average in the national population is 7.5 times. The premium for tertiary has, however, been increasing at a higher rate for Africans than is seen in the full population. Africans with tertiary earned five times the salary of those with primary in 1994 and seven times the salary in 2010. The overall national tertiary premium increased from 7.5 times the salary of those with primary in 1994 to 8.5-9 times the salary by 2010. Second, the declines in the matric and incomplete secondary premiums are more subdued within the African only sample than are seen in the national population estimates. The matric (incomplete secondary) premium declines by 16% (25%) for Africans compared to the decline of 30% (40%) for the overall population. Thus here again we see convergence in the matric and incomplete secondary premiums between the African only and national population estimates.

Figure 5. Relative earnings for incomplete secondary, matric and tertiary compared to primary – African only versus national population– all years

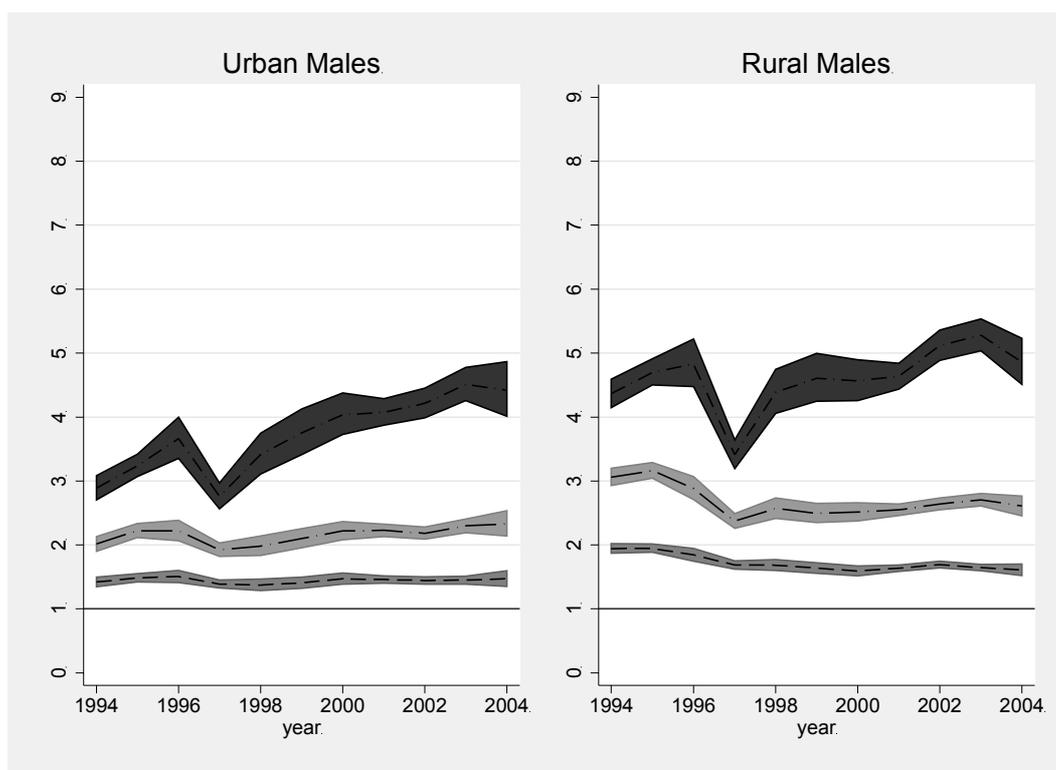


Note: The figure presents converted wage regression coefficients (and confidence bands) on incomplete secondary (bottom line), matric (middle line) and tertiary (top line) relative to primary or less from a linear regression of log earnings, controlling for a quadratic in age for Africans only (left panel) and the national population (right panel). Converted coefficients = $\exp(b)$. All regressions weighted using the cross entropy weights.

Urban and rural

In Figure 6 we revert once more to the comparison of tertiary, matric and incomplete secondary premiums relative to primary. Here they are presented for urban and rural areas separately. Note the urban/rural classification was only available until 2004. Premiums to higher levels of education are higher in rural versus urban areas, reflecting in part the scarcity of higher levels of education in rural areas. However, the premium to tertiary, while having increased in both areas, has increased more substantially in urban areas than rural areas. In 1994, the male tertiary premium in urban areas was about 3 times the earnings of males with primary in urban areas, while in rural areas the comparable estimate was 4.5 to 1. In 2004 the male tertiary premium in urban areas was 4.5:1, while the rural male tertiary premium was 5:1. The matric and incomplete secondary premiums are also higher in rural areas than urban area. Here again we see convergence over time. The male-rural-matric premium decreased from 3 to 2.5 between 1994 and 2004, while the male-urban-matric premium increased from 2 to 2.1. By 2004, the incomplete secondary premium was similar in urban and rural areas.

Figure 6. Relative earnings for incomplete secondary, matric and tertiary compared to primary – Urban and rural subsamples



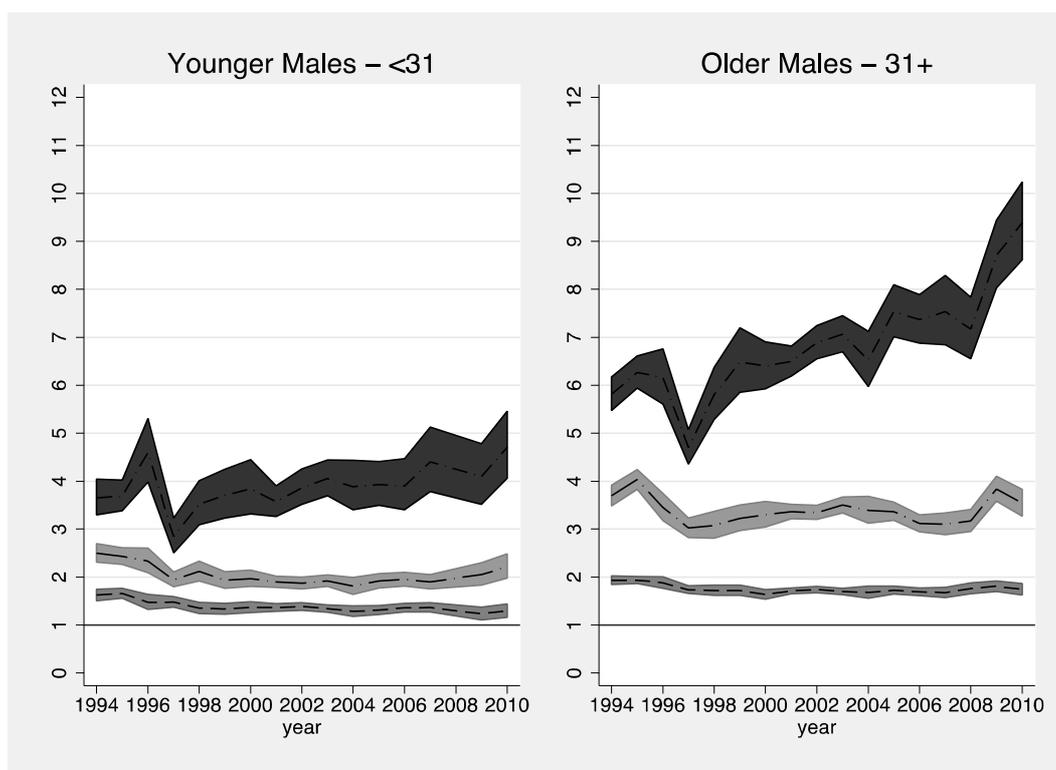
Note: The figure presents the converted wage regression coefficients (and 95% confidence bands) on incomplete secondary (bottom line), matric (middle line) and tertiary (top line) relative to primary or less from a linear regression of log earnings, controlling for a quadratic in age and coloured, Indian and white indicators. Separate regression results presented for urban and rural areas. Sample restricted to males. Converted coefficients = $\exp(b)$. All regressions weighted using the cross entropy weights.

Young (15-30) and older (31-64) workers

Figure 7 compares the relative earnings premiums to education level within younger and older worker subsamples. Restricting the sample to younger workers (15-30) highlights that younger cohorts have not been the recipients of the increased premium over the period evident in Figure 3. The tertiary premium among younger male workers has only increased marginally over the period to about 4 times the salary of someone with primary in 2010. In the older workers group, the tertiary premium increased from about 6:1 in 1994 to 9:1 in 2010.

In each case we see higher premiums to tertiary for subsamples with lower average educational attainment. This is to be expected. However, some interesting points are evident from Figures 5-7. First, the African only and national population tertiary premiums are converging. This signals that Africans with tertiary education are, in particular, increasingly valued in the labour market. Second, there is a convergence in returns to education category between urban and rural areas. Educational attainment has increased primarily at the bottom tail of the distribution in rural areas while improvements have been in the middle in urban areas. Hence the average educational attainment of the rural reference category has increased and converged towards the urban reference category mean. Finally, the divergence in the tertiary premium between the young and older worker samples over time indicates that the South African labour market increasingly values not only workers with tertiary but more experienced workers with tertiary education. In fact the tertiary premium within the younger subsample does not increase much over the period examined. This is evidence that the age profile of those with tertiary is much steeper than for workers with lower levels of education.

Figure 7. Relative earnings for incomplete secondary, matric and tertiary compared to primary – younger (15-30) and older workers (31-64)

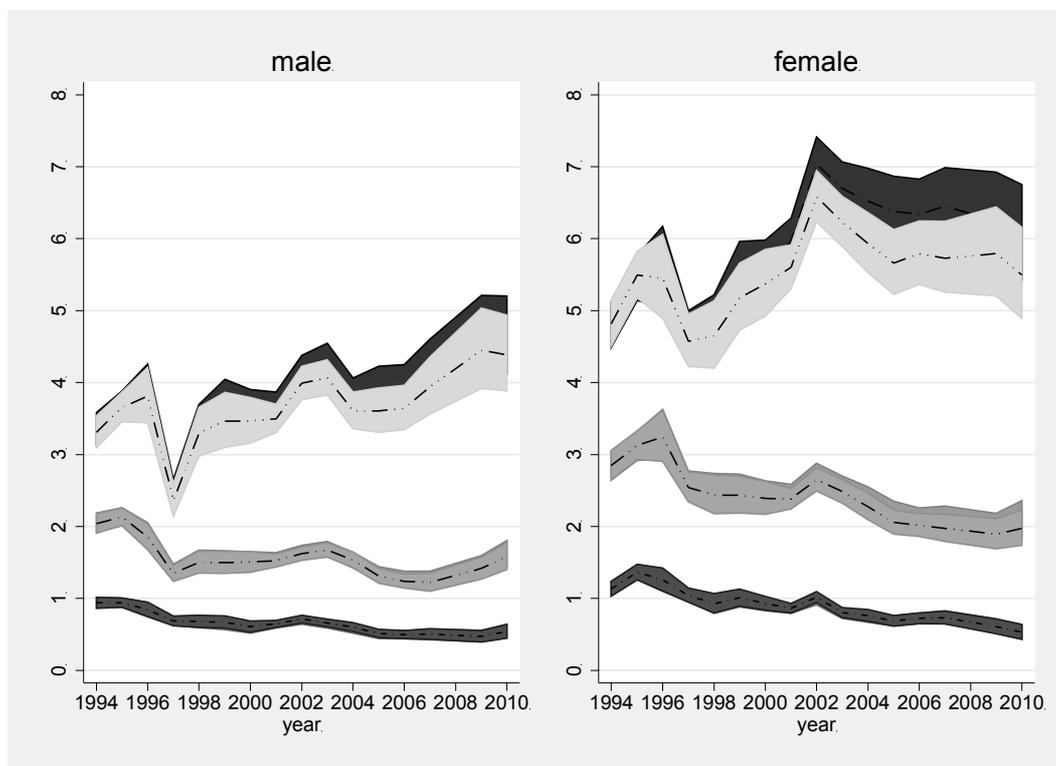


Note: The figure presents the converted wage regression coefficients (and 95% confidence bands) on incomplete secondary (bottom line), matric (middle line) and tertiary (top line) relative to primary or less from a linear regression of log earnings, controlling for a quadratic in age and coloured, Indian and white indicators. Separate regression results presented for young respondents (15-30) and older respondents (31-64). Sample restricted to males. Converted coefficients = $\exp(b)$. All regressions weighted using the cross entropy weights.

2.3.3. Testing for presence of sample selection bias

Figure 8 presents the relative monthly earnings that those with incomplete secondary, matric and tertiary received when compared to workers with primary, once an attempt is made to account for selection into employment as described in Section 2.2. In addition, we include for comparison purposes, the converted education coefficients from equation 1 (in the specification including race dummies). The estimates controlling for selection into employment are in the foreground and the original estimates are in the background.

It is clear from the figure that selection, as measured using the instruments set out in Section 2.2, is only present for the tertiary group. This is in line with the results in Figure 4, which show that only those with tertiary have an increased probability of employment once the full model is specified. We see that once selection is taken into account the tertiary premiums are smaller, falling more in the later years. These estimates are not however significantly different from the original wage estimates as is evident in the overlapping confidence intervals.

Figure 8. Comparison of education level coefficients from Heckman selection model and control model

Note: Converted wage regression coefficients (and 95% confidence bands) on incomplete secondary (bottom lines), matric (middle lines) and tertiary (top lines) relative to primary or less are presented. The estimates in the foreground are from a heckman maximum likelihood selection regression. The earnings regression includes a quadratic in age and coloured, Indian and white indicators. The employment equation includes, in addition, marital status, number of children under 7, between 7 and 18 and number of working adults in the household. The estimates in the background replicate those from the earnings regression in including a full set of controls. Converted coefficients = $\exp(b)$. All regressions weighted using the cross entropy weights.

3. Concluding remarks

We use 17 years of national household survey data and a simple and consistent model specification to illustrate the relationship between educational attainment and labour market earnings and employment probability over time. We show large differences in the size and trend of the premium to different levels of education. Both the earnings and employment probability premiums to tertiary are high and have increased over the period. Matric and incomplete secondary earnings premiums remain positive, but have stayed at constant levels, if not decreased, over the period. There is no employment premium to matric and incomplete secondary education once controls for population group, household composition and urban residence are taken into account.

These trends are consistent with shifts in the demand and supply of labour by education level. South Africa is experiencing skills biased growth. The share of employment in manufacturing and agriculture has declined and the share in trade, finance and construction increased. Educational output is not aligned with this demand for skills. There has been an increase in labour force participants with education below tertiary, but only a small increase in the share of individuals with tertiary education. Thus the increase in the tertiary premium appears driven by an increased demand for skills that are not being readily produced in the South African education system. We saw however, that it is particularly older individuals and African tertiary individuals that are increasingly valued in terms of earnings in the labour market.

That being said, the aim of the paper is to describe the levels and trends in the education category premiums over a period that experienced many changes. The levels and trends of premiums to different education levels raises many questions about the underlying mechanisms driving these difference, but no effort is placed on identifying these in this paper. One particular area for future work is the fact that racial controls account for a large proportion of the premium. This brings up questions of how much of the premiums are a function of observable differences between population groups such as quality of education and connectedness to the labour market and how much can be attributed to discrimination.

BIBLIOGRAPHY

- Branson, N., Garlick, J., Lam, D. and M. Leibbrandt (2012), "Education and Inequality: The South African Case", *SALDRU Working Papers*, No 75.
- Branson, N. and M. Wittenberg (2007), "The measurement of employment status in South Africa using cohort analysis, 1994-2004", *South African Journal of Economics*, Vol.75(2), pp. 313-326.
- Burger, R.P. and D. Yu, (2006), "Wage trends in post-Apartheid South Africa: Constructing an Earnings Series from Household Survey Data", *Stellenbosch Economic Working Papers*.
- Burger R. and R. Jafta (2006), "Returns to Race: Labour Market Discrimination in Post-Apartheid South Africa", Department of Economics, University of Stellenbosch.
- Burger, C. and S. van Der Berg (2011), "Modelling cognitive skills, ability and school quality to explain labour market earnings differentials", *Stellenbosch Economic Working Papers*, 08/11.
- Burger, R. and D. von Fintel (2009), "Determining the Causes of the Rising South African Unemployment Rate: An Age, Period and Generational Analysis", *Stellenbosch Working Papers*, No. 158.
- Casale, D. and D. Posel (2002), "The Continued Feminisation of the Labour Force in South Africa: An Analysis of Recent Data and Trends", *South African Journal of Economics*, Vol. 70(1), pp. 156-184.
- Chamberlain, D. and S. van der Berg (2002), "Earnings Functions, Labour Market Discrimination and Quality of Education in South Africa", *Stellenbosch Economic Working Papers*, 2/2002.
- Kerr, A. and M. Wittenberg (2012), "The Impact of Changes in Statistics South Africa's Enumeration Practice on Average Household Size", Presentation to CSAE Conference, March 2012, available at https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=CSAE2012andpaper_id=565.
- Kingdon, G. and J. Knight (2008), "Unemployment: South Africa's Achilles Heel" in *South African Economic Policy under Democracy*. Aron, J., Kahn, B. and G. Kingdon (eds). Oxford: Oxford University Press.
- Leibbrandt, M., Woolard, I., Finn A. and J. Argent (2010), "Trends in South African Income Distribution and Poverty since the Fall of Apartheid", *Social, Employment and Migration Working Papers*, No. 101, OECD Publishing, available at <http://dx.doi.org/10.1787/5kmms0t7p1ms-en>.
- Leibbrandt, M., Woolard, I., McEwen, H. and C. Koep (2010), "Employment and Inequality Outcomes in South Africa", *OECD Working Papers*.
- Rospabé, S. (2002), "How did labour market racial discrimination evolve after the end of apartheid? An analysis of the evolution of employment, occupational and wage discrimination in South Africa between 1993 and 1999", *The South African Journal of Economics*. Vol. 70(1), pp. 185-217.
- Spaull, N. (2012), "Poverty and Privilege: Primary School Inequality in South Africa", *International Journal of Educational Development*, pp. 1-12.
- Wooldridge, J. (2002), "Econometric analysis of cross section and panel data", *Economic Analysis of Cross Sections and Panel Data*, MIT Press, England.

APPENDIX 1

SURVEY SAMPLE INFORMATION AND VARIABLE DEFINITIONS

Survey sample information:

The data consists of 17 years of national household survey data from 1994 to 2010; the October Household Surveys (OHS) between 1994 and 1999, the Labour Force Surveys (LFSs) between 2000 and 2007 and the General Household Surveys (GHSs) between 2008 and 2010. The OHS was a broad-based survey instrument that aimed to collect a wide range of socioeconomic information. From 2000 the LFS, a rotating panel collected twice each year, became the primary source of labour market information with the GHS forming its counterpart from 2002 onwards. In 2008, the biannual LFS was replaced by the quarterly LFS and earnings are currently not available in these data. The GHS has less information on the labour market (although it does include earnings) and has more detailed education and other social variables than the LFS. With the exception of OHS 1996, OHS 1998 and February LFS 2000 (which were smaller), each survey sampling frame was approximately 30 000 households. We use a set of weights (cross entropy weights) that are calibrated to consistent demographic and geographic estimates over time (Branson, 2010) to improve the consistency of representation of the surveys over time. Our analysis focuses on men and women aged 15-64.

Highest level of education was coded into years of education as presented in Table A1. The variable differed slightly across years, most notably becoming more detailed over time. Table A1 clarifies which categories were grouped together. The ‘years of education’ variable was then grouped into four distinct categories, namely primary (0-7 years), incomplete secondary (8-11 years), matric (12 years) and tertiary (13-16 years) and formed the basis of the dummy variables included in the wage and employment regressions.

Labour force variables:

The board definition of labour force participation was used in this paper. People were classified as economically active if they indicated that they wanted work even if they had not actively sought work in the last seven days. Employed included any respondent who was working in wage employment or self-employment, full or part-time. The unemployed are those labour force participants not classified as employed.

The earnings variable used represents total monthly earnings before tax and deductions including overtime, allowances and bonuses. Wage earnings and self-employed earnings are not disaggregated in the regression analysis. Where point estimates were not given, the midpoint of the bracket amount was used. Brackets have remained the same since 1995 and individuals could choose between reporting monthly, weekly or annual bracketed or exact amounts. The brackets in 1994 were odd with the first bracket large and the brackets thereafter very detailed.

Tables 1-3 and Figure 1 present real earnings in 1994 rands. Earnings in each year are deflated to 1994 rands using Statistics South Africa’s monthly consumer price index (P0141). A trimmed earnings variable was calculated and the sensitivity of estimates to outliers assessed. Burger and Yu (2006) suggest that a 2000 real value of ZAR 83 336 presents an appropriate cut-off for outliers. We adjusted this cut-off to 1994 values and used the result – ZAR 56 475 - as a cut-off in each year. Table A2 presents the number of outliers by survey year.

*Education variables:***Table A1. Classification of highest education level into years of education**

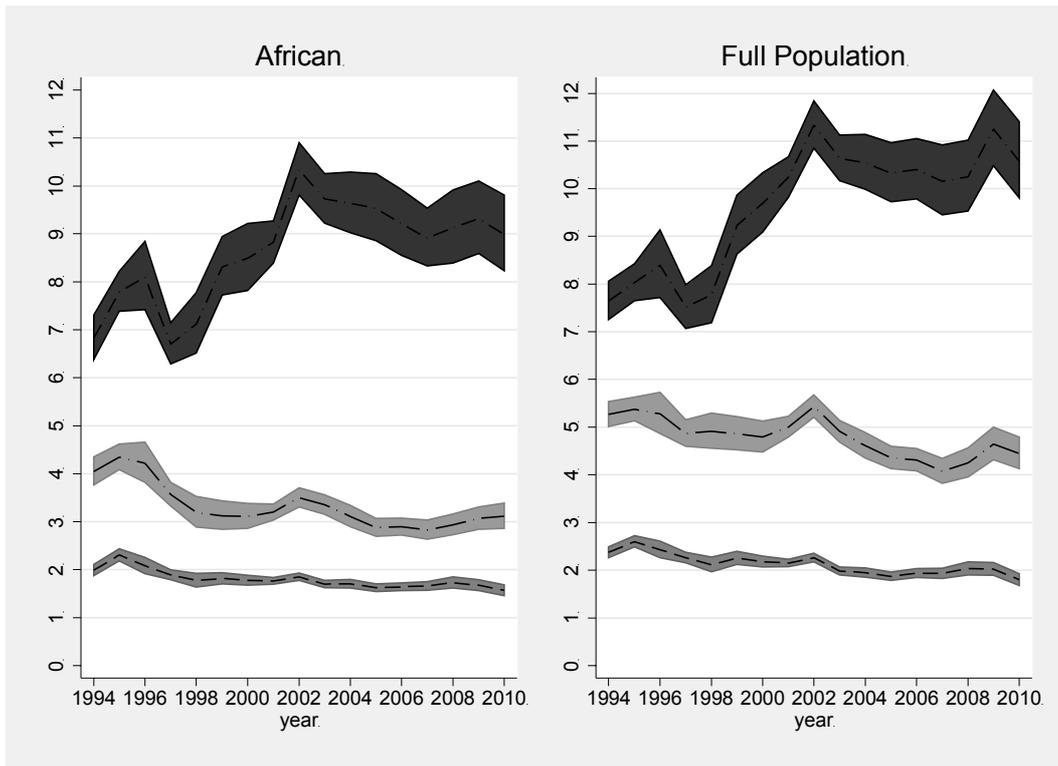
Years of education	OHS 1994-1995	OHS 1996	OHS 1997-LFS 2007	GHS 2008	GHS 2009 - 2010
0	No schooling	No schooling	No schooling	No schooling	No schooling
0	Grade 0	Grade 0	Grade 0	Grade R/0	Grade R/0
1	Grade 1/grade2/grade 3				
1		Grade 1	Grade 1	Sub A/Grade 1	Grade 1/ Sub A/Class 1
2		Grade 2	Grade 2	Sub B/Grade 2	Grade 2 / Sub B/Class 2
3		Grade 3	Grade 3	Grade 3/Standard 1	Grade 3/Standard 1/ABET 1(Kha Ri Gude; Sanli)
4	Grade 4	Grade 4	Grade 4	Grade 4/Standard 2	Grade 4/ Standard 2
5	Grade 5	Grade 5	Grade 5	Grade 5/Standard 3	Grade 5/ Standard 3/ ABET 2
6	Grade 6	Grade 6	Grade 6	Grade 6/Standard 4	Grade 6/Standard 4
7	Grade 7	Grade 7	Grade 7	Grade 7/Standard 5	Grade 7/Standard 5/ ABET 3
8	Grade 8	Grade 8	Grade 8	Grade 8/Standard 6/Form 1	Grade 8/Standard 6/Form 1
9	Grade 9	Grade 9	Grade 9	Grade 9/Standard 7/Form 2	Grade 9/Standard 7/Form 2/ ABET 4
10	Grade 10/NTC I	Grade 10/NTC I	Grade 10/NTC I	Grade 10/Standard 8/Form 3	Grade 10/ Standard 8/ Form 3
10				NTC I	NTC 1/ N1/NC (V) Level 2
11	Grade 11/NTC II	Grade 11/NTC II	Grade 11/NTC II	Grade 11/Standard 9/Form 4	Grade 11/ Standard 9/ Form 4
11				NTC II	NTC 2/ N2/ NC (V) Level 3
11	Certificate or diploma with less than	Certificate or diploma with less than	Certificate or diploma with less than	Certificate with less than grade 12	Certificate with less than Grade 12
11				Diploma with less than grade 12/STD	Diploma with less than Grade 12/Std
12	Grade 12/NTC III	Grade 12/NTC III	Grade 12/NTC III	Grade 12/Standard 10/Form 5/Matric	Grade 12/Standard 10/Form 5/Matric (No Exemption)
12				NTC III	Grade 12/Standard 10/Form 5/Matric (Exemption *)
12					NTC 3/ N3/NC (V)/Level 4
12					N4/NTC 4
13	Certificate or diploma with grade 12	Certificate or diploma with grade 12	Certificate or diploma with grade 12	Certificate with grade 12/STD 10	N5/NTC 5
13				Diploma with grade 12/STD 10	N6/NTC 6
13					Certificate with Grade 12/Std 10
13					Diploma with Grade 12/Std 10
13					Higher Diploma (Technikon/University)
15			Undergraduate degree	Bachelor's Degree	Post Higher Diploma (Technikon/University)
15					Bachelors Degree
16	Degree (undergrad or postgrad)	Degree (undergrad or postgrad)	Post-graduate degree/diploma	Bachelor's Degree and Diploma	Bachelors Degree and post-graduate
16				Honours Degree	Honours Degree
16				Higher Degree (Masters, Doctorate)	Higher degree (Masters, Doctorate)
.	Other	Other	Other	Other	Other (specify in the box below)
.				Do not know	Do not know
.				Unspecified	Unspecified

Table A2. Number of observations classified as outliers (real earnings in 1994 rands > 56 475) by survey year

<u>Survey year</u>	<u>Number of outliers</u>
1994	78
1995	110
1996	20
1997	42
1998	34
1999	47
2000	26
2001	9
2002	3
2003	3
2004	2
2005	6
2006	5
2007	8
2008	6
2009	9
2010	5
<u>Total</u>	<u>413</u>

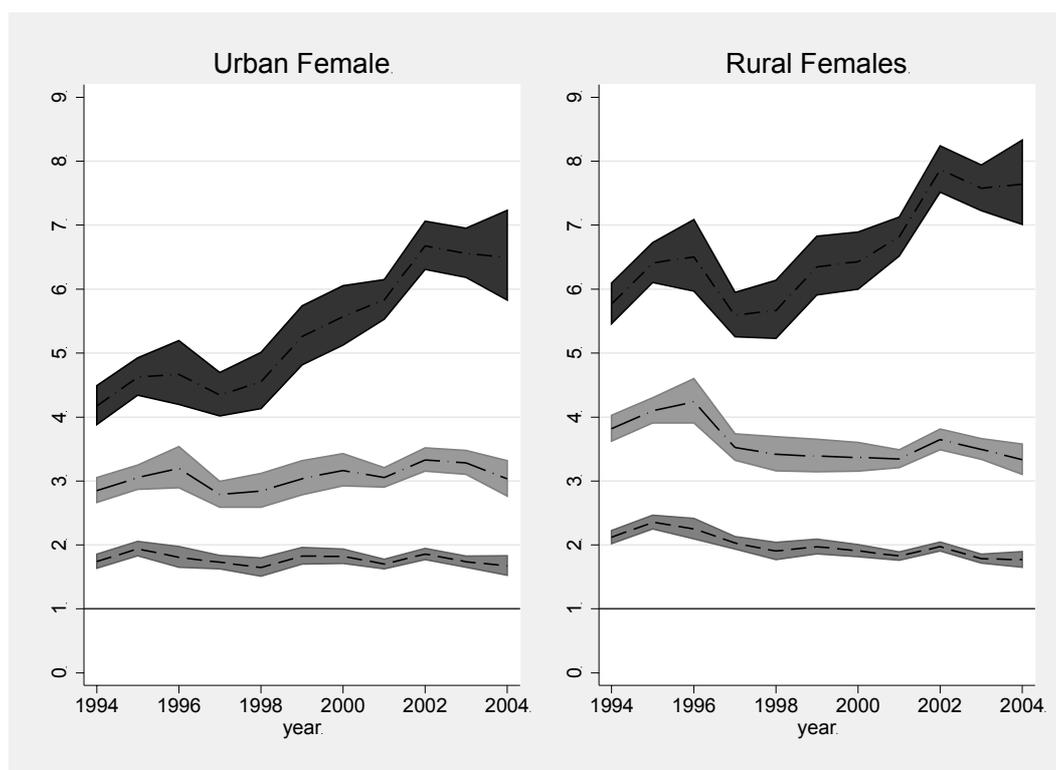
APPENDIX 2
OTHER FIGURES

Figure B1. Relative earnings for matric and tertiary compared to incomplete secondary or less – African only versus national population – females only, all years



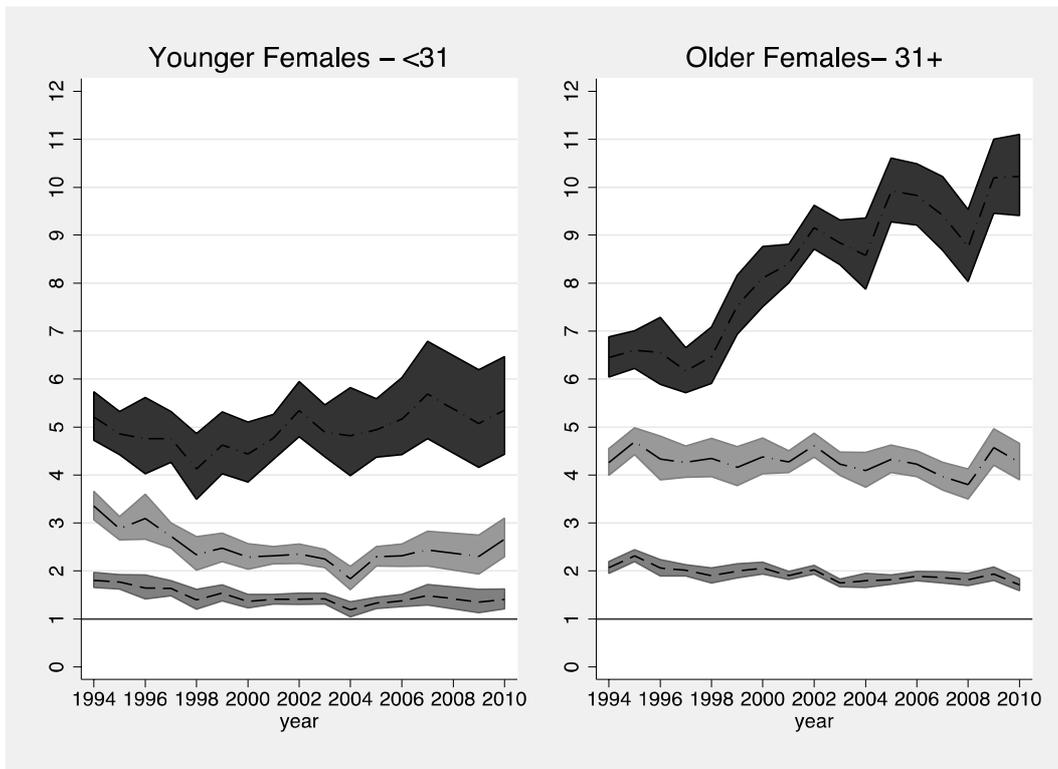
Note: The figure presents converted wage regression coefficients (and confidence bands) on incomplete secondary (bottom line), matric (middle line) and tertiary (top line) relative to primary or less from a linear regression of log earnings, controlling for a quadratic in age for Africans only (left panel) and the national population (right panel). Converted coefficients = $\exp(b)$. All regressions weighted using the cross entropy weights. Females only.

Figure B2. Relative earnings for incomplete secondary, matric and tertiary compared to primary – Urban and rural subsamples



Note: The figure presents the converted wage regression coefficients (and 95% confidence bands) on incomplete secondary (bottom line), matric (middle line) and tertiary (top line) relative to primary or less from a linear regression of log earnings, controlling for a quadratic in age and coloured, Indian and white indicators. Separate regression results presented for urban and rural areas. Sample restricted to females. Converted coefficients = $\exp(b)$. All regressions weighted using the cross entropy weights.

Figure B3. Relative earnings for incomplete secondary, matric and tertiary compared to primary – younger (15-30) and older workers (31-64)



Note: The figure presents the converted wage regression coefficients (and 95% confidence bands) on incomplete secondary (bottom line), matric (middle line) and tertiary (top line) relative to primary or less from a linear regression of log earnings, controlling for a quadratic in age and coloured, Indian and white indicators. Separate regression results presented for young respondents (15-30) and older respondents (31-64). Sample restricted to females. Converted coefficients = $\exp(b)$. All regressions weighted using the cross entropy weights.

WORKING PAPERS

The full series of Economics Department Working Papers can be consulted at www.oecd.org/eco/workingpapers/

- 1021. *Education quality and labour market outcomes in South Africa*
(January 2013) by Nicola Branson and Murray Leibbrandt
- 1020. *Do policies that reduce unemployment raise its volatility? Evidence from OECD countries*
(January 2013) by Alain de Serres and Fabrice Murtin
- 1019. *Slovakia: A catching up euro area member in and out of the crisis*
(January 2013) by Jarko Fidrmuc, Caroline Klein, Robert Price and Andreas Wörgötter
- 1018. *Improving the fiscal framework to enhance growth in an era of fiscal consolidation in Slovakia*
(January 2013) by Caroline Klein, Robert Price and Andreas Wörgötter
- 1017. *Investing efficiently in education and active labour market policies in Slovakia*
(January 2013) by Caroline Klein
- 1016. *The performance of road transport infrastructure and its links to policies*
(January 2013) by Henrik Braconier, Mauro Pisu and Debra Bloch
- 1015. *The US labour market recovery following the great recession*
(January 2013) by Wendy Dunn
- 1014. *Why do Russian firms use fixed-term and agency work contracts?*
(December 2012) by Larisa Smirnykh and Andreas Wörgötter
- 1013. *The Equity implications of fiscal consolidation*
(December 2012) by Lukasz Rawdanowicz, Eckhard Wurzel and Ane Kathrine Christensen
- 1012. *The Dutch labour market: preparing for the future*
(December 2012) by Mathijs Gerritsen and Jens Høj
- 1011. *Reforming policies for the business sector to harvest the benefits of globalisation in the Netherlands*
(December 2012) by Mathijs Gerritsen and Jens Høj
- 1010. *Health care reform and long-term care in the Netherlands*
(December 2012) by Erik Schut, Stéphane Sorbe and Jens Høj
- 1009. *Enhancing the inclusiveness of the labour market in Belgium*
(December 2012) by Jens Høj
- 1008. *Reducing poverty in Estonia through activation and better targeting*
(December 2012) by Sarah Flèche and Artur Radziwill
- 1007. *Matching skills and jobs in Estonia*
(December 2012) by Lilas Demmou

1006. *Debt and macroeconomic stability: An overview of the literature and some empirics*
(December 2012) by Douglas Sutherland and Peter Hoeller
1005. *Debt and macroeconomic stability: Debt and the business cycle*
(December 2012) by Volker Ziemann
1004. *Debt and macroeconomic stability: Case studies*
(December 2012) by Rossana Merola
1003. *Debt and macroeconomic stability*
(December 2012) by Douglas Sutherland, Peter Hoeller, Rossana Merola and Volker Ziemann
1002. *Reducing greenhouse gas emissions in a cost effective way in Switzerland.*
(December 2012) by Anita Wölfl and Patrizio Sicari
1001. *Strengthening innovation in the United States*
(November 2012) by David Carey, Christopher Hill and Brian Kahin
1000. *Long-term growth scenarios*
(January 2013) by Åsa Johansson, Yvan Guillemette, Fabrice Murin, David Turner, Giuseppe Nicoletti, Christine de la Maisonneuve, Philip Bagnoli, Guillaume Bousquet and Francesca Spinelli
999. *Selected aspects of household savings in Germany – evidence from micro-data*
(November 2012) by Christina Kolerus, Isabell Koske and Felix Hüfner
998. *Improving the tax system in Indonesia*
(November 2012) by Jens Arnold
997. *Unleashing business innovation in Canada*
(November 2012) by Alexandra Bibbee
996. *Public policy and resource allocation: evidence from firms in OECD countries*
(October 2012) by Dan Andrews and Federico Cingano
995. *Promoting SME development in Indonesia*
(October 2012) by Annabelle Mourougane
994. *Portugal: Rebalancing the economy and returning to growth through job creation and better capital allocation.*
(October 2012) by Álvaro Pina and Ildeberta Abreu
993. *Public debt, economic growth and nonlinear effects: Myth or reality?*
(October 2012) by Balázs Égert
992. *Choosing the pace of fiscal consolidation*
(September 2012) by Lukasz Rawdanowicz
991. *Tertiary education developing skills for innovation and long-term growth in Canada*
(September 2012) by Calista Cheung, Yvan Guillemette and Shahrzad Mobasher-Fard